Journal of Protein Chemistry, Vol. 15, No. 1, 1996

Oligophosphopeptides of Varied Structural Complexity Derived from the Egg Phosphoprotein, Phosvitin

Antonis Goulas,[1] E. L. Triplett,[1] and George Taborsky[1,2]

Received May 24, 1995

Phosvitins are the principal phosphoproteins in the eggs of oviparous vertebrates. They have an exceptionally high serine content and most, or even all, of the serine residues are esterified to phosphate. The phosphorylated residues tend to occur in uninterrupted runs of as many as 28 phosphoserines (as in *Xenopus* phosvitin). This unique structural feature gives phosvitins extraordinary properties and can be expected to play a key role in phosvitin function. For example, the concentration of phosphate groups provides for numerous highly efficient metal-binding sites in clusters. The mode of binding had been shown to be affected by the size of the protein and the degree to which serine residues are phosphorylated. For structure–function studies of phosvitins (and other polyphosphoproteins), phosphopeptides of differentiated structural complexity are desirable. Such model peptides were produced in this work by limited proteolysis of chicken phosvitin, and oligophosphopeptides of widely varying sizes, phosphoserine content, and sequence were purified and characterized. These include phosvitin segments containing one, two, or several oligophosphoserine runs, corresponding to segments of the N-terminal, C-terminal, and core sequence of the protein.

KEY WORDS: Phosvitin; proteolysis; phosphopeptide fragments; phosphoprotein models.

1. INTRODUCTION

Past studies carried out in our laboratory examined some of the exceedingly highly phosphorylated proteins of the egg yolk, the phosvitins. These proteins contain a very large number of phosphoserine residues, which are ordered within the primary structure as highly distinctive oligophosphoserine segments. Our past studies included investigations of the manner in which phosvitins accommodate iron, exploiting the chelating properties of the phosphoryl groups (Grogan and Taborsky, 1986, 1987). We realized that such studies could be greatly aided by the availability of relatively simple oligophosphopeptides, each representing only some part of the structural complexity of the intact protein molecule in which several distinctive structural motifs coexist. Such peptides could also be used as models in structure–function studies of other polyphosphoproteins.

Proteolytic fragmentation seemed to provide an obvious approach to the production of simple phosvitin "models." However, phosvitin is particularly resistant to proteolytic action *in vitro*. This could be attributed to the unusual primary structure of phosvitins containing long oligophosphoserine blocks uninterrupted by other residues. These runs can be as long as 14 phosphoserines, as in the chicken (Byrne *et al.*, 1984), or even 28 residues, as in *Xenopus* (Gerber-Huber *et al.*, 1987). Such runs represent an extraordinary concentration of negative charge that may well interfere with proper enzyme–substrate complex formation.

[1] Department of Biological Sciences, Division of Molecular, Cellular and Developmental Biology, University of California, Santa Barbara, California 93106.

[2] To whom correspondence should be addressed.

0277-8033/96/0100-0001$09.50/0

Early proteolytic experiments with the chicken protein indicated (Mecham and Olcott, 1949; Belitz, 1963) that no more than two or three peptide bonds per molecule of phosvitin are readily susceptible to action by trypsin, for example, in spite of the fact that arginine and lysine together constitute 26 residues per mole of M_r 35,000. Peptic digestion was found to be similarly restricted. While Belitz (1965, 1966) was able to isolate and characterize proteolytic fragments of various lengths and compositions, the mapping of these fragments with respect to the primary structure of the intact protein was not possible since sequence data were unknown at the time. Also, the nonuniform character of the material used in those studies would have handicapped any mapping efforts.

It seemed to us that with the availability of highly purified phosvitins of known sequence, attempting to overcome the sluggishness of phosvitin proteolysis would be justified. We report here the generation and isolation, in good yield, of peptic, chymotryptic, and tryptic fragments from highly purified preparations of the major chicken phosvitin (Clark, 1970). These fragments were characterized in terms of their amino acid composition and phosphate content. On the basis of these characterizations, we were able to assign these fragments to specific segments of the known sequence of the protein (Byrne *et al.*, 1984) by means of a "best fit" of the data. Also, we could confirm some of these assignments by direct sequence analyses of the peptides.

2. MATERIALS AND METHODS

2.1. Phosvitin Preparations

Chicken phosvitin was isolated according to the method of Joubert and Cook (1958). The major component was purified as described by Clark (1970) and was used in all experiments.

2.2. Proteolytic Digestion

Conditions previously described by Belitz (1963) guided our initial experiments. Reactions were conducted in a jacketed vessel connected to a water bath for temperature control. The pH stability of the reaction mixtures during tryptic and chymotryptic digestions was ensured by means of continuous titration with NaOH utilizing a Radiometer pH-stat.

Trypsin, chymotrypsin, and pepsin were purchased from Worthington as crystalline powders: pepsin and trypsin were twice-crystallized preparations, PM and TRL grade, respectively; chymotrypsin was a low-molecular-weight, contaminant-free, CDS-grade preparation. Prior to use, all three enzymes were dissolved in 0.001 N HCl according to the manufacturer's specifications (0.5 mg/ml pepsin; 10 mg/ml chymotrypsin or trypsin). Enzymic activities, assayed according to protocols provided by Worthington, were 650, 54, and 120 units/mg for pepsin, chymotrypsin, and trypsin, respectively. These values corresponded to 20%, 89%, and 61%, respectively, of the original enzymic activities reported by Worthington. The activity losses are attributable to long periods of refrigerated storage and were deemed acceptable for these studies. Initial enzyme:substrate ratios were typically on the order of 1:100 (w:w) for all three enzymes, although higher ratios were used occasionally, as indicated below.

Phosvitin was dissolved in 3 ml of water to a final concentration of 15 mg/ml and the pH of the solution was adjusted to 8.0 with NaOH prior to the addition of either trypsin or chymotrypsin, or to pH 2.5 with HCl prior to the addition of pepsin. Temperature was maintained at 26°C, 30°C, and 37°C, for tryptic, peptic, and chymotryptic digestions, respectively. Reactions were allowed to proceed for 24 hr. Tryptic and chymotryptic digestions were arrested by acidification to pH 5.3 with HCl. Peptic digestions were arrested by neutralization with NaOH.

2.3. Size Fractionation and Purification of Proteolytic Products

Enzymic digests were size-fractionated on a G-50 Sephadex (Pharmacia) column (2 × 80 cm) equilibrated with 0.5 M NaCl at room temperature. Elution profiles were observed in terms of absorption measurements at 220 nm (LKB Ultrospec 4050 spectrophotometer). In addition, the eluted fractions were assayed for alkali-labile (protein-bound) phosphate with automatic flow analysis performed with a Technicon system (McCollum and Taborsky, 1983). Fractions representing individual peaks were combined and

desalted either on a G-15 Sephadex column equilibrated with water or by repeated dialysis (Spectrapor tubing) against large volumes of water at room temperature. In some cases, as indicated, additional size fractionation was carried out on a G-75 Sephadex column equilibrated with 10 mM NH_4HCO_3. Most gel filtration products were further purified through anion exchange chromatography performed on a DE-51 (Whatman) column equilibrated with 50 mM Tris-HCl, pH 8.0, and eluted with a NaCl gradient (0–0.5 M). Purified fragments were stored at −20°C as lyophilized powders.

2.4. Characterization of Proteolytic Fragments

Samples for amino acid analysis were hydrolyzed and analyzed, using the ninhydrin method, according to Spackman *et al.*, (1958). Filtered hydrolyzates were analyzed using a Varian Model 5500 HPLC system, equipped with a 0.4 × 15 cm Micropak "hydrolyzate column" operated with the manufacturer's buffer solutions. Data were collected and analyzed using an Apple IIe computer with Analab hardware and the Chromatochart program of Interactive Microwave. Corrections were applied for losses of serine (25%) (Allerton and Perlmann, 1965) and threonine (10%) (Allerton and Perlmann, 1965; Hirs *et al.*, 1954).

Individual proteolytic fragments were assigned to segments of the known phosvitin primary structure on the basis of amino acid compositional data, searching for a "best fit" between composition and sequence. First, an estimate of the size of the fragment was obtained by attribution of the smallest amino acid residue percent value (>0·5) to a single residue. The amino acid composition of the fragment was then compared to that of the intact protein. Since the sequence of chicken phosvitin (shown in Fig. 5) is heavily biased with respect to the location of several amino acids in the sequence, the relative enrichment or deficiency of such amino acids in the fragment under consideration could be used to identify the portion of the phosvitin sequence from which the fragment had to be derived. For example, a composition rich in alanine is indicative of N-terminal origin since alanine residues are found exclusively in the N-terminal part of the protein. The same is true for threonine-rich fragments. Conversely, the presence of histidine, leucine, or isoleucine points to a C-terminal origin. Amino acids that occur only once in the sequence—such as phenylalanine, methionine, tryptophan, or tyrosine—help further limit the proposed origin of a fragment to within specific bounds. Once a tentative sequence assignment is made in this manner, compositional values—in terms of residue percent—are calculated for the putative segment and compared with the actual analytical data obtained for the fragment. The tentative sequence of the fragment is refined through the addition or subtraction of individual amino acid residues at the termini of the sequence in order to produce a "best fit" between postulated sequence and analytical composition. A "best fit" was defined as one for which the number of "deviant" amino acids would be zero or very small. Amino acids were considered to be deviant if the ratio of analytical-to-predicted residue percent values was smaller than 0.8 or greater than 1.2.

Confirmation of "best-fit" sequence assignments was sought through direct, partial N-terminal sequence analysis of some of the proteolytic fragments. The analysis was accomplished by automatic sequencing performed with a Porton 2090E Model sequencer, using the Edman method, with trimethylamine as the coupling buffer (method no. 40, in the instrument manufacturer's manual). All chemicals used in the sequence determinations were obtained from Porton Instruments.

3. RESULTS

3.1. Peptic Digestion

Peptic proteolysis of phosvitin produced three identifiable fragments in good yield. The largest of these fragments, produced with a 1:100 enzyme:substrate ratio (w:w), was eluted just past the excluded volume (V_0) from a G-50 column (Fig. 1A, peak 1). Comparison of its amino acid composition with that of the intact protein led to the assignment of this fragment to the segment contained within residues Asn 44 and Ile 215 of the phosvitin sequence (Fig. 5). Material produced with higher relative enzyme concentrations [E:S = 6:100 and 1:10 (w:w)] was eluted at the same position (Fig. 1B, C, peak 1). However, this material provided a better fit with a slightly shorter phosvitin segment, Asn 44–Leu 193 (Table I), indicative of a more complete cleavage near the C-terminus. This was reinforced by the finding that the yield of the smaller fragment obtained from the G-50 column

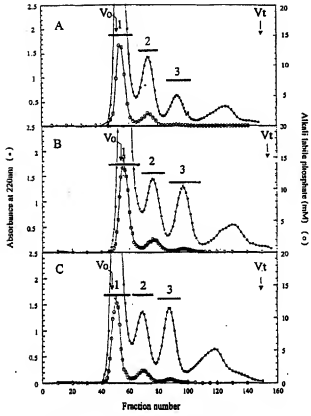


Fig. 1. Gel filtration of peptic digests of the major component of chicken phosvitin, obtained with three different E:S ratios (w:w): (A) 1:100; (B) 6:100; (C) 1:10. In each case, the reaction mixture (3 ml), containing 45 mg of pepsin-digested phosvitin, was loaded onto a Sephadex G-50 column (2 × 80 cm) equilibrated with 0.5 M NaCl. The column was eluted with the NaCl solution at the rate of 0.6 ml/min. Individual fractions of about 1 ml were collected. Fractions were analyzed for absorbance at 220 nm and alkali-labile phosphate concentration (mM) as shown. Peak fractions 1, 2, and 3 were pooled as designated. V_0 and V_t denote the excluded and the total column volumes, respectively.

(Fig. 1A, peak 3) increased, in terms of absorbance as well as the amount of alkali-labile phosphate, as the enzyme concentration was raised. This fragment corresponds to the C-terminal segment Glu 194–Glu 214 (Table I). The composition of the intermediate-size fragment (Fig. 1A–C, peak 2) remained relatively unchanged as the relative concentration of enzyme was raised. This fragment was fitted best to the sequence limited by Gly 4 and Glu 41 (Table I).

The digest produced with the highest relative enzyme concentration (E:S = 1:10; Fig. 1C) yielded peaks 1, 2, and 3 (obtained with the G-50 column) which represented 81.4%, 14.9%, and 3.5% of the total alkali-labile phosphate. It is noteworthy that the values for peaks 1 and 2 are in reasonably close agreement with the calculated percentage values of total phosvitin-bound phosphate that would be associated with the segments to which these fragments were fitted best: 79.0% for Asn 44–Leu 193 and 13.7% for Gly 4–Glu 41. These calculations assume that all serines and one threonine of phosvitin are phosphorylated (Grogan *et al.*, 1990). The phosphate in peak 3, 3.5%, is only about half of the expected value, 7.3%, for the segment Glu 194–Glu 214. This may be attributed to the slow rate of cleavage of the peptide bond (Leu 193–Glu 194) whose proteolysis must produce the C-terminal fragment Glu 194–Glu 214.

Unequivocal confirmation of the sequence assignments based on our "best fit" analysis was obtained by direct N-terminal sequence determination for the short C-terminal fragment and the longer N-terminal fragment. No such confirmation could be obtained for the large fragment assigned to the Asn 44–Leu 193 segment: the sequencer results were equivocal, presumably because of some heterogeneity of the fragment and because of the poor response of phosphoserine in the Edman reaction.

For the fragment providing a best fit to the segment Glu 194–Glu 214, a clear identification of three N-terminal residues (Glu–Asp–Asp) left little doubt regarding the true identity of the fragment. No other residues appeared as "contaminants," suggesting no significant heterogeneity of the purified fragment. Following the Glu–Asp–Asp sequence, the quality of the sequencer product deteriorated sharply, but not unexpectedly, given that the known phosvitin sequence following Asp 196 consists of a hexapeptide composed of phosphoserines, a very poorly reactive, decomposing product of the Edman reaction.

The N-terminal sequence Gly 4–Glu 41 was also firmly confirmed by direct sequencing. In this case, the sequencer yielded unequivocal data for an octapeptide sequence (Gly–Thr–Glu–Pro–Asp–Ala–Lys–Thr) which uniquely agrees with phosvitin residues 4–11. At position 12, where the sequencer data deteriorate sharply, the known phosvitin sequence begins a phosphoserine-rich segment.

3.2. Tryptic Digestion

The G-50 elution profile of a typical tryptic digest is shown in Fig. 2. Of the observed peaks,

Table I. Analytical Amino Acid Composition of Chicken Phosvitin Fragments (residue/mol)[a]

Fragment[b]	P-1		P-2		P-3		T-1		T-2		C-1		C-2	
"Best fit"	Asn 44–Leu 193		Gly 4–Glu 41		Glu 194–Glu 214		Gln 49–Arg 212		Ala 1–Arg 35		Ala 50–Trp 210		Gly 4–Gln 49	
Lys	**11.77**	(11)	**2.88**	(3)	**1.20**	(1)	**10.67**	(11)	**1.11**	(1)	**10.67**	(11)	**3.74**	(4)
His	**11.32**	(12)	+	(0)	**1.03**	(1)	**10.56**	(12)	—	(0)	**10.68**	(12)	+	(0)
Arg	**9.00**	(9)	1.37	(1)	**0.96**	(1)	**8.10**	(10)	2.09	(1)	**8.64**	(9)	**1.09**	(1)
Asx	**6.72**	(8)	**3.42**	(3)	**1.84**	(2)	**8.72**	(8)	**2.30**	(2)	**8.48**	(8)	**4.65**	(5)
Thr	—	(0)	—[c]	(4)	—[c]	(0)	—	(0)	2.40	(4)	—	(0)	**3.60**	(4)
Ser	**97.02**	(98)	**22.20**[c]	(16)	**9.99**[c]	(9)	**108.07**	(107)	**16.32**	(16)	**109.14**	(107)	19.84	(16)
Glx	**2.52**	(3)	**1.94**	(2)	**1.74**	(2)	**2.76**	(3)	**2.26**	(2)	**2.18**	(2)	**5.28**	(6)
Pro	—	(0)	(+)	(3)	—	(0)	—	(0)	1.42	(2)	—	(0)	+	(3)
Gly	3.93	(3).	**1.14**	(1)	**0.93**	(1)	**4.32**	(4)	1.48	(1)	**3.60**	(3)	1.31	(1)
Ala	2.50	(2)	**3.36**	(4)	+	(0)	**2.04**	(2)	**4.85**	(5)	2.46	(2)	**4.40**	(4)
Val	**2.04**	(2)	—	(0)	**0.87**	(1)	**2.00**	(2)	—	(0)	**1.96**	(2)	**0.98**	(1)
Met	—	(0)	0.58	(1)	—	(0)	—	(0)	—	(0)	—	(0)	0.78	(1)
Ile	—	(0)	—	(0)	1.29	(1)	0.79	(1)	—	(0)	0.77	(1)	—	(0)
Leu	**2.32**	(2)	—	(0)	**0.93**	(1)	**2.97**	(3)	—	(0)	**2.79**	(3)	—	(0)
Tyr	—	(0)	—	(0)	—	(0)	—	(0)	—	(0)	—	(0)	—	(0)
Phe	—	(0)	—	(0)	—	(0)	—	(0)	?	(1)	—	(0)	—	(0)
Trp	ND	(0)	ND	(0)	ND	(1)	ND	(1)	ND	(0)	ND	(1)	ND	(0)
Total residues		(150)		(38)		(21)		(164)		(35)		(161)		(46)

[a] Analytical data are given in boldface whenever the agreement with the value for the "best-fitting" sequence (given in parentheses) was 20% or less (see Section 2). A + designates amino acids detected in the sample but absent from the "best fitting" segment. A (+) designates an amino acid found to be present but not quantitatively determined. A ? represents an amino acid present in the "best-fitting" segment but not detected in the sample. ND signifies that the amino acid (tryptophan) was not determined.

[b] Fragment designations identify the proteolytic enzyme that produced the fragment (P, pepsin; T, trypsin; C, chymotrypsin) and the peak number assigned in the respective Sephadex G-50 elution profiles (Figs 1, 2, and 4).

[c] Value for serine includes both serine and threonine because of incomplete resolution in the analytical chromatogram.

only the two major ones—peak 1, at V_0, and peak 2, immediately following—contained alkali-labile phosphate. Peak 1 was further resolved upon passage through a DE-51 anion exchange column (Fig. 3). Of the several peaks obtained, only one contained protein-bound phosphate. Its amino acid, composition fitted that of a 164-amino acid residue-long segment beginning with Gln 49 and ending with Arg 212 of the phosvitin sequence (Fig. 5; Table I). Peak 2 represented an obviously

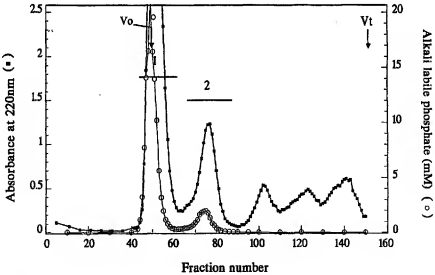


Fig. 2. Gel filtration of a tryptic digest [E:S = 1:100 (w:w)] of the major component of chicken phosvitin. For details, see the legend of Fig. 1.

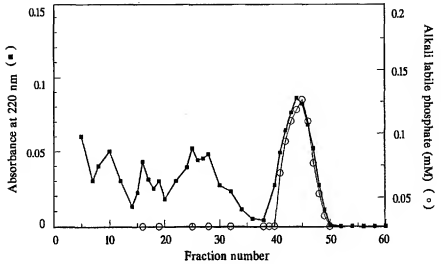


Fig. 3. Anion exchange chromatography of peak 1 material recovered from the tryptic phosvitin digest (Fig. 2). A total of 3.4 mg of desalted, lyophilized material was dissolved in about 1 ml of 50 mM Tris-HCl, pH 8.0, and applied to a Whatman DE-51 column (2.2 × 26.5 cm) equilibrated with the same buffer. The column was eluted with a NaCl gradient produced with a two-chamber system containing 400 ml of the starting (50 mM Tris-HCl, pH 8.0) or limiting (50 mM Tris-HCl, pH 8.0, 0.5 M NaCl) buffer. The gradient was started at fraction number 15. Fractions of about 5 ml were collected at the rate of about 1.5 ml/min. All phosphate-containing fractions were pooled as indicated. Only a single phosphate-containing peak was eluted at about fraction number 40, corresponding to a NaCl concentration of about 1 M.

heterogeneous collection of fragments, since its composition could not be related to a single phosvitin segment without allowing for a large number of "deviant" amino acids (see Section 2). The situation could not be improved even when an additional size fractionation (with G-75 Sephadex; not shown) was introduced into the purification scheme. The fact, however, that the peak 2 material is rich in alanine and threonine and deficient in methionine indicates that this material must be derived primarily from within segment Ala 1–Arg 35 (Fig. 5; Table I).

Of the total protein-bound phosphate that was eluted from the G-50 column, 86.9% was associated with peak 1 and 13.1% with peak 2. These values are in good agreement with the theoretical percentages of the total phosvitin-bound phosphate estimated for the segment assigned to the major components of either peak 1 (86.3%) or peak 2 (13.7%).

Raising the relative trypsin concentration to E:S = 1:10 (w:w) did not appear to produce additional fragmentation, as judged from the G-50 elution profile (not shown). This remained essentially identical with the one shown in Fig. 2. Also, small variations in pH (between 7.5 and 9) or an increase in temperature (to 30°C) produced no significant variations in the elution pattern. Extensive neutralization of the negative charge on phosvitin, by saturating all of the phosphate metal-binding sites with iron, produced no change in the elution profile.

3.3. Chymotryptic Digestion

Fractionation of a chymotryptic digest [E:S = 1:100 (w/w)] on a G-50 Sephadex column produced a profile similar to the one obtained after tryptic digestion. Two alkali-labile, phosphate-containing components were clearly resolved (Fig. 4; peaks 1 and 2). This pattern persisted if the enzyme concentration was raised (E:S = 1:10; results not shown). Both the larger component (peak 1) and the smaller one (peak 2) represented apparently homogeneous material since neither one was resolved any further upon additional anion exchange chromatography on a DE-51 column (results not shown). The amino acid composition of the fragment which was eluted as a peak 1 was fitted best to the phosvitin segment Ala 50–Trp 210

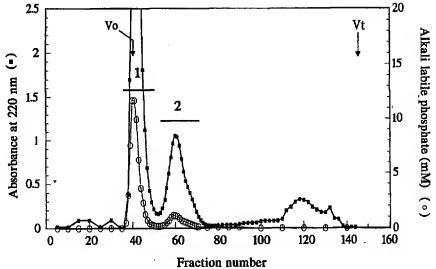


Fig. 4. Gel filtration of a chymotryptic digest [E:S = 1:100 (w:w)] of the major component of chicken phosvitin. For details, see the legend of Fig. 1.

(Table I), which is almost identical with the segment assigned to the large tryptic fragment (see Section 3.2, above). The smaller fragment (peak 2) was assigned to a segment near the N-terminus: Gly 4–Gln 49 (Table I).

Percent alkali-labile phosphate of the total, recovered as constituent of the larger or smaller fragment, was 84.9% and 15.1%, respectively. The theoretical values for the respective segments are 86.3% and 13.7%.

4. DISCUSSION

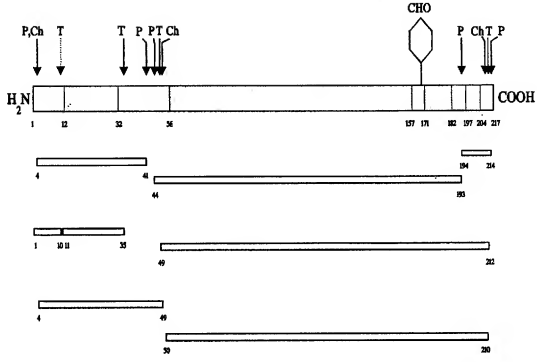
All results of this study are summarized in Fig. 5. The figure shows the results obtained with all three enzymes used in this study: trypsin, chymotrypsin, and pepsin.

Pepsin, but not trypsin or chymotrypsin, produced a small fragment, apparently by accessing a small opening (His 183–Asp 196) between adjacent phosphoserine blocks near phosvitin's C-terminus. This fragment contains a single, eight-residue-long oligophosphoserine run. In contrast to the N-terminal segments (see below), which were produced in essentially quantitative yields, the proteolytic release of the C-terminal fragment was no more than half complete even when the E:S ratio was as high as 1:10 (w:w). A reason for this sluggishness may be suggested as follows. At pH 2—the pH of the peptic digestion mixture—the negative charges of phosvitin are extensively suppressed and would not be expected to interfere with effective interaction with enzyme. However, the protein is known to undergo a conformational transition toward a β-type structure at low pH (Taborsky, 1968). This could result in making a cleavage site less accessible. Also, the oligosaccharide attached at Asn 169 (Byrne *et al.*, 1984), being relatively near the postulated cleavage site, might hinder enzymic access.

All three enzymes used effected the release of N-terminal fragments by cleaving within a 23-residue-long, phosphoserine-free sequence limited by phosvitin residues Pro 33 and Ala 55. Each of these fragments contains two oligophosphoserine runs that account for 16 of phosvitin's 123 phosphoserines. They also contain all four of the variably phosphorylated threonine residues of chicken phosvitin (Grogan *et al.*, 1990).

Following the loss of N- and C-terminal segments and some "trimming" at the two ends, the core of the protein remained largely intact upon digestion with any of the three proteolytic enzymes. This is not surprising, given the unusual constitution of the phosvitin core, which is made up entirely of oligophosphoserine runs interrupted only by basic (Lys or Arg) amino acid residues in a relatively few places (Fig. 5). In view of the fact that the negative charge of the phosphate group has been shown to render neighboring peptide bonds insensitive to tryptic action (Theodoropoulos *et al.*,

A



B

```
          10         20         30         40         50
          .          .          .          .          .
AEFGTEPDAKTSSSSSASSTATSSSSSSASSPNRKKPMDEEENDQVKQA
          60         70         80         90        100
          .          .          .          .          .
RNKDASSSSRSSKSSNSSKRSSSKSSNSSKRSSSSSSSSSSSSRSSSSS
         110        120        130        140        150
          .          .          .          .          .
SSSSNSKSSSSSSKSSSSSSRSRSSSKSSSSSSSSSSSSSSSKSSSSRSSS
         160        170        180        190        200
          .          .          .          .          .
SSSKSSSHHSHSHHSGHLNGSSSSSSSSRSVSHHSHEHHSGHLEDDSSSS
         210
          .
SSSSVLSKIWGRHEIYQ
```

Fig. 5. (A) A simplified representation of phosvitin's primary structure is shown at the top. Arrows point to presumed sites of peptic (P), tryptic (T), and chymotryptic (Ch) attack. Numbers indicate the position of amino acid residues that define the terminal residues of phosphoserine-rich segments (cross-hatched boxes). CHO denotes the position of the oligosaccharide component of phosvitin. Underneath the phosvitin structure, a schematic representation of peptic, tryptic, and chymotryptic fragments (open bars) of chicken phosvitin is given. The fragments are assigned to particular segments of the phosvitin amino acid sequence to which they were best fitted (see text). Numbers denoting the beginning and end of each fragment correspond to the respective terminal amino acid positions. (B) The amino acid sequence of the major chicken phosvitin, as reported by Byrne *et al.* (1984).

1959*a*, *b*), the peptide bonds adjacent to lysine or arginine residues are evidently inaccessible to proteases in phosvitin. Also, the heterogeneity of the tryptic N-terminal fragment may be explained in terms of the varied access to trypsin of the peptide bond between Lys 10 and Thr 11, given that the latter could be variably phosphorylated (Grogan *et al.*, 1990).

An attempt we made to neutralize the charge effect by saturating the protein with iron did not enhance fragmentation. While undoubtedly the cationic metal bound at anionic phosphate binding sites would appreciably suppress the net negative charge, the formation of a properly constituted enzyme:substrate complex may be blocked in this case not by charge, but by bulk—given that phosvitin-bound ion may form polynuclear complexes (Gray, 1971).

Several of the proteolytic phosvitin fragments we prepared, purified, and characterized are uniquely assignable to specific segments of the phosvitin sequence. Each fragment contains particular numbers of phosphoserines in particular molecular environments. It is especially noteworthy that these fragments range in complexity from the relatively simple near-C-terminal segment 194–214 containing just a single octapeptide run of phosphoserines, through several near-N-terminal segments (1–35, 4–41, 4–49), all of them containing two identical decapeptide sequences, each with a hexaphosphoserine run, to several core fragments (44–193, 49–212, 50–210) with nearly 70 residue percent of phosphoserines arranged in runs of up to 14 residues in length.

These fragments represent a potentially valuable set of phosphopeptides, as they can serve as well-defined models of varied complexity for structure–function studies of phosvitins and other polyphosphoproteins. A ^{31}P–NMR study of such fragments has already been carried out in our laboratory, yielding fresh insights into structural aspects of phosvitin (J. Grogan, manuscript in preparation).

ACKNOWLEDGMENTS

This study was supported in part by a Research Grant from the U.S. National Institutes of Health (GM 32750-06). We are grateful for the expert technical contributions of Kathrin McCollum to various aspects of this study, particularly to the purification of the proteolytic phosvitin fragments, their analyses for alkali-labile phosphate content, and their characterization in terms of partial amino acid sequence data. We thank also Dr. James Grogan and Dr. Kathy Kozak for some of the fractionations of proteolytic digests conducted on a preparatory scale.

REFERENCES

Allerton, S. E., and Perlmann, G. E. (1965). *J. Biol. Chem.* **240,** 3892–3898.
Belitz, H.-D. (1963). *Z. Lebensm. Unters. Forsch.* **119,** 381–389.
Belitz, H.-D. (1965). *Z. Lebensm. Unters. Forsch.* **127,** 341–352.
Belitz, H.-D. (1966). *Z. Lebensm. Unters. Forsch.* **130,** 152–157.
Byrne, B. M., van het Schip, A. D., Van De Klundert, J. A. M., Arnberg, A. C., Gruber, M., and Ab, G. (1984). *Biochemistry* **23,** 4275–4279.
Clark, R. C. (1970). *Biochem. J.* **118,** 537–542.
Gerber-Huber, S., Nardelli, D., Haeflinger, J.-A., Cooper, D. N., Givel, F., Germont, J.-E., Engel, J., Green, N. M., and Wahli, W. (1987). *Nucleic Acid Res.* **15,** 4737–4759.
Gray, H. B. (1971). *Adv. Chem.* **100,** 368–389.
Grogan, J., and Taborsky, G. (1986). *J. Inorg. Biochem.* **26,** 237–246.
Grogan, J., and Taborsky, G. (1987). *J. Inorg. Biochem.* **29,** 33–47.
Grogan, J., Shirazi, A., and Taborsky, G. (1990). *Comp. Biochem. Physiol.* **96B,** 655–669.
Hirs, C. H. W., Stein, W. H., and Moore, S. (1954). *J. Biol. Chem.* **211,** 941–950.
Joubert, F. J., and Cook, W. H. (1958). *Can. J. Biochem. Physiol.* **36,** 399–408.
McCollum, K., and Taborsky, G. (1983). *Anal. Biochem.* **130,** 311–320.
Mecham, D. K., and Olcott, H. S. (1949). *J. Am. Chem. Soc.* **71,** 3670–3679.
Spackman, D. H., Stein, W. H., and Moore, S., (1958). *Anal. Chem.* **30,** 1190–1206.
Taborsky, G. (1968). *J. Biol. Chem.* **243,** 6014–6020.
Theodoropoulos, D., Bennich, H., Folsch, G., and Mellander, O. (1959*a*). *Nature* **184,** 187–188.
Theodoropoulos, D., Bennich, H., and Mellander, O. (1959*b*). *Nature* **184,** 270–271.